**Information Retrieval**

1. **June 2014**

**Ex 1 [ranks 4]** Construct the Cartesian Tree for the array A = [4, 2, 6, 8, 1], and specify the property that relates this tree with Range Minimum Queries over A.

**Ex 2 [ranks 5+5]** Given a posting list describe the use of skip pointers:

* Comment on the best skip-positioning in terms of time access, whenever all items have the same probability to be accessed.
* Comment on the best positioning of skips in terms of time access, whenever each item has its own probability to be accessed.

**Ex 3 [points 4+3]** Given the following matrix of pair-wise similarity between 5 items



Show the next cluster formed by the agglomerative clustering algorithm based on MIN and MAX similarity-functions, given that we have already formed the following clusters {(I1,I2), (I3,I4), (I5)}

**Ex 4 [points 4]** Design a cuckoo hash over two tables of size m=7, each, and hash functions h1(k) = k mod m, h2(k) = 3\*k mod m. Show the content of the two tables by inserting the sequence of keys S={1, 5, 8, 3, 12, 10, 11, 13, 9} if possible.

**Ex 5 [points 2+3]** Given a graph of three pages, and directed edges (1,3) (2,3) (3,2).

1. Specify the formula of HITS on this graph, and indicate how can be computed a(3) and h(3) as a function of the other node’s values.
2. Assuming that the initial values for a() and h() are equal to 1, make one step of HITS calculation.